IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

PATENT

In re application of: Khaleel

Attorney Docket No.: NWISP049

Application No.: 10/775,974

Examiner: LAU, TUNG S.

Filed: February 9, 2004

Group: 2863

Title: HISTOGRAM PERFORMANCE COUNTERS FOR USE IN TRANSACTION LATENCY ANALYSIS

CERTIFICATE OF MAILING
I hereby certify that this correspondence is being deposited with the U.S. Postal Service with sufficient postage as first-class mail on November 22, 2005 in an envelope addressed to the Commissioner for Patents, P.O. Box 1450 Alexandria, VA 22313-1450.

Signed: ______________________
Mia Mitchell-Haynes

PRE-APPEAL BRIEF REQUEST FOR REVIEW

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Applicant requests review of the final rejection in the above-identified application. No amendments are being filed with this request.

This request is being filed with a Notice of Appeal.

The review is requested for the reasons stated on the attached sheets.

Remarks begin on page 2 of this paper.

REMARKS

The Examiner Mischaracterized the Teachings of the Primary Reference

In the final office action dated September 23, 2005, the Examiner finally rejected claims 1-33 and 35-46 of the present application under 35 U.S.C. 102(b) as being anticipated by U.S. Patent Publication No. 2002/0049824 A1 (Wilson). The Examiner has consistently mischaracterized the teachings of Wilson, and has inappropriately applied these mischaracterizations to multiple limitations in the claims of the present application.

Simply stated, the main issue to be appealed is as follows: Wilson does not teach or suggest either the latency or histogram counters recited in the claims of the present application.

Wilson describes a multi-processor computer architecture having a distributed shared memory. The system includes a plurality of uniform memory access (UMA) cells 100 interconnected by an interconnect 20. Each UMA cell includes a processor 102 and a memory 104 which includes a memory controller 106. Each memory controller 106 includes a so-called "cache of history counter" (CofHC) 108. Each entry in CofHC 108 represents a page (indicated by page address field 112) in the associated memory 104 and includes a plurality of counters (e.g., 114-120 of Fig. 2). Each counter for a given entry tracks a different metric for the corresponding memory page. For example, each of counters 114, 115, and 116 corresponds to one of the other UMA's (e.g., 200, 300, etc.) in the system, and tracks the number of times the processor in the corresponding UMA accesses the memory page represented by that entry. Migration counter 118 is incremented upon each migration of the page represented by the CofHC from one cache to another. Write counter 120 is incremented for each write to the page. By tracking these metrics, the system can determine (i.e., using the decision tree 250 of Fig. 3) whether there is a more efficient memory location for a given page in memory. See paragraphs [0023]-[0030].

As is well known in the field of computing, the term "latency" refers to a duration of time associated with execution of a given type of operation, e.g., a memory access. As stated in the Webopedia (**http://www.webopedia.com/TERM/L/latency.html**), latency refers to "the period of time that one component in a system is spinning its wheels waiting for another component. Latency, therefore, is wasted time. For example, in accessing data on a disk, latency is defined as the time it takes to position the proper sector under the read/write head."

Significantly, none of the counters described by Wilson determines or tracks transaction latency or any period of time. That is, as described above, Wilson's counters track the number of times specific types of events occur in the system, e.g., how many times a cached memory page is accessed by specific processors in the system, so that it can then determine whether the page

should be moved or copied to another cache. None of the quantities counted represent a period of time, i.e., a latency. In fact, there is nothing in Wilson which tracks latency, particularly for individual system transactions, e.g., memory or I/O accesses. Therefore, none of Wilson's counters can be characterized as a latency counter.

By contrast, claim 1 of the present invention recites "a latency counter operable to generate a latency count for each of selected" memory transactions. That is, the recited latency counter maintains a count which is representative of the length of time (e.g., as measured in clock cycles) required to complete all or a portion of a particular memory or I/O transaction.

Claim 1 also recites "a plurality of histogram counters," each of which is "operable to count selected ones of the latency counts corresponding to an associated latency range." That is, each histogram counter tracks the number of transaction latencies (as determined by the latency counter) which fall within a particular latency range. For example, one histogram counter might be responsible for counting latency counts which fall within the range 600-700 clock cycles while another would be responsible for counting latency counts which fall within the range 700-800 clock cycles, and so on. So, not only does the invention recited in claim 1 include the tracking of latencies for specific memory transactions, it also includes a second level of counting (a counting of the latency counts), i.e., where each of these transaction latencies fits with respect to multiple ranges within an overall time window.

An exemplary implementation is described in the present application from page 32, line 7 to page 33, line 21, with reference to Fig. 14. When a memory or I/O transaction begins, the Transaction Start signal is asserted, and when the transaction ends, the Transaction End signal is asserted. Between these two events, a latency counter counts the number of clock cycles. This maps to the latency counter and the latency count recited in claim 1. The resulting latency count is then placed in a "bucket" which corresponds to a latency range which includes that latency count, i.e., a histogram counter corresponding to the latency range is incremented. These "buckets" map to the histogram counter recited in claim 1. The present invention thus allows detailed latency data, e.g., a distribution of transaction latencies, to be accumulated during system operation.

Wilson neither describes nor suggests the determination or tracking of any type of latency as recited in claim 1 of the present application. Wilson does note that the information collected by the CoHC counters may be used to determine whether a particular memory page should be replicated or migrated to a different memory location which, in turn, may result in improvements in system latency (see paragraph [0030]). However, Wilson does not discuss how or even whether such latency improvements are determined. Latency measurements are clearly not performed or tracked using the CoHC counters or any of the other components of Wilson's system.

And because Wilson does not teach or suggest the tracking of individual transaction latencies, it also does not teach or suggest the tracking of how many such transaction latencies correspond to particular latency ranges. That is, Wilson also fails to show the use of histogram counters to count the latency counts as recited in claim 1.

The Examiner has continued to refer to paragraphs [0029] and [0030] of Wilson as teaching the tracking of latency despite the clear lack of such teaching in the reference. The Examiner also referred to the term "history counters" in the Abstract of Wilson as corresponding to the "histogram counters" recited in the claims of the present application, despite the fact that this term clearly refers to the same counters.

The "history counters" of Wilson, i.e., counters 114-120 in the cache of history counters (CofHC) 108, track the "history" of access and migrations for a particular page in memory. See [0027]-[0029]. By contrast, and as is well known in mathematics and engineering, a "histogram" is "a bar graph of a frequency distribution in which the widths of the bars are proportional to the classes into which the variable has been divided and the heights of the bars are proportional to the class frequencies." See The American Heritage College Dictionary, Third Edition. In the claims of the present application, the variable is transaction latency and the classes are the latency ranges, the number of latencies falling into each range being the frequency of that class.

The Applicants' representative attempted to discuss the matter in a telephone interview with the Examiner and it became clear that the Examiner simply did not understand either the reference or the claimed invention. Neither did the Examiner acknowledge that he was using the same teaching to reject two different limitations of the claims, or appear to understand the difference between the terms "history" and "histogram."

Because Wilson fails to teach or suggest a latency counter which is operable to generate latency counts for transactions involving memory, and because Wilson also fails to teach or suggest a plurality of histogram counters for counting the latency counts for corresponding latency ranges, it is respectfully submitted that the rejection of the claims over Wilson should be withdrawn.

Respectfully submitted,
BEYER WEAVER & THOMAS, LLP

Joseph M. Villeneuve
Attorney of Record
Reg. 37,460

P.O. Box 70250
Oakland, CA 94612-0250
510-663-1100